Published in final edited form as:

Science. 2010 December 10; 330(6010): 1551–1557. doi:10.1126/science.1195271.

The Major Genetic Determinants of HIV-1 Control Affect HLA Class I Peptide Presentation

The International HIV Controllers Study[*,†], **Writing team**, **Florencia Pereyra**[1,2,‡], **Xiaoming Jia**[3,‡], **Paul J. McLaren**[4,5,‡], **Amalio Telenti**[6], **Paul I.W. de Bakker**[4,5,7,8,†][co-chair], **Bruce D. Walker**[1,9,‡][co-chair], **Analysis team**, **Xiaoming Jia**[3], **Paul J. McLaren**[4,5][project leaders], **Stephan Ripke**[4,10], **Chanson J. Brumme**[1], **Sara L. Pulit**[4,5], **Amalio Telenti**[6], **Mary Carrington**[1,11], **Carl M. Kadie**[12], **Jonathan M. Carlson**[13], **David Heckerman**[13], **Paul I.W. de Bakker**[4,5,7,8,†][chair], **Study design**, **Florencia Pereyra**[1,2], **Paul I.W. de Bakker**[4,5,7,8,†], **Robert R. Graham**[14], **Robert M. Plenge**[4,15], **Steven G. Deeks**[16], **Bruce D. Walker**[1,9,†][chair], **SNP genotyping, HLA typing, and sample management**, **Lauren Gianniny**[4], **Gabriel Crawford**[4], **Jordan Sullivan**[4], **Elena Gonzalez**[4], **Leela Davies**[4], **Amy Camargo**[4], **Jamie M. Moore**[4], **Nicole Beattie**[4], **Supriya Gupta**[4], **Andrew Crenshaw**[4], **Noël P. Burtt**[4], **Candace Guiducci**[4], **Namrata Gupta**[4], **Mary Carrington**[1,11], **Xiaojiang Gao**[11], **Ying Qi**[11], **Yuko Yuki**[11], **HIV controllers recruitment and sample management**, **Florencia Pereyra**[1,2][project leader], **Alicja Piechocka-Trocha**[1], **Emily Cutrell**[1], **Rachel Rosenberg**[1], **Kristin L. Moss**[1], **Paul Lemay**[1], **Jessica O'Leary**[1], **Todd Schaefer**[1], **Pranshu Verma**[1], **Ildiko Toth**[1], **Brian Block**[1], **Brett Baker**[1], **Alissa Rothchild**[1], **Jeffrey Lian**[1], **Jacqueline Proudfoot**[1], **Donna Marie L. Alvino**[1], **Seanna Vine**[1], **Marylyn M. Addo**[1], **Todd M. Allen**[1], **Marcus Altfeld**[1], **Matthew R. Henn**[4], **Sylvie Le Gall**[1], **Hendrik Streeck**[1], **Bruce D. Walker**[1,9,†][chair], **AIDS Clinical Trials Group**, **David W. Haas**[17], **Daniel R. Kuritzkes**[2], **Gregory K. Robbins**[18], **Robert W. Shafer**[19], **Roy M. Gulick**[20], **Cecilia M. Shikuma**[21], **Richard Haubrich**[22], **Sharon Riddler**[23], **Paul E. Sax**[2], **Eric S. Daar**[24], **Heather J. Ribaudo**[25], **HIV controllers referral team**, **Brian Agan**[26], **Shanu Agarwal**[27], **Richard L. Ahern**[18], **Brady L. Allen**[28], **Sherly Altidor**[29], **Eric L. Altschuler**[30], **Sujata Ambardar**[31], **Kathryn Anastos**[32], **Ben Anderson**[33], **Val Anderson**[34], **Ushan Andrady**[34], **Diana Antoniskis**[35], **David Bangsberg**[1,18], **Daniel Barbaro**[36], **William Barrie**[37], **J. Bartczak**[38], **Simon Barton**[39], **Patricia Basden**[40], **Nesli Basgoz**[18], **Suzane Bazner**[1], **Nicholaos C. Bellos**[41], **Anne M. Benson**[40], **Judith Berger**[42], **Nicole F. Bernard**[43], **Annette M. Bernard**[44], **Christopher Birch**[1], **Stanley J. Bodner**[45], **Robert K. Bolan**[46], **Emilie T. Boudreaux**[47], **Meg Bradley**[1], **James F. Braun**[48], **Jon E. Brndjar**[49], **Stephen J. Brown**[50], **Katherine Brown**[51], **Sheldon T. Brown**[52], **Jedidiah Burack**[53], **Larry M. Bush**[54], **Virginia Cafaro**[55], **Omobolaji Campbell**[18], **John Campbell**[56], **Robert H. Carlson**[57], **J. Kevin Carmichael**[58], **Kathleen K. Casey**[59], **Chris Cavacuiti**[60], **Gregory Celestin**[61], **Steven T. Chambers**[62], **Nancy Chez**[63], **Lisa M. Chirch**[64], **Paul J. Cimoch**[65], **Daniel Cohen**[66], **Lillian E. Cohn**[67], **Brian Conway**[68], **David A. Cooper**[69], **Brian Cornelson**[60], **David T. Cox**[70], **Michael V. Cristofano**[71], **George Cuchural Jr.**[72], **Julie L. Czartoski**[73], **Joseph M. Dahman**[74], **Jennifer S. Daly**[75], **Benjamin T. Davis**[18], **Kristine Davis**[76], **Sheila M. Davod**[18], **Steven G. Deeks**[16], **Edwin DeJesus**[77], **Craig A. Dietz**[78], **Eleanor Dunham**[64], **Michael E. Dunn**[79], **Todd B. Ellerin**[80], **Joseph J. Eron**[81], **John J.W. Fangman**[82], **Claire E. Farel**[2], **Helen Ferlazzo**[83], **Sarah Fidler**[84], **Anita Fleenor-Ford**[85], **Renee Frankel**[86], **Kenneth A. Freedberg**[18], **Neel K. French**[87], **Jonathan D. Fuchs**[88], **Jon D. Fuller**[89], **Jonna Gaberman**[90], **Joel E. Gallant**[91], **Rajesh T. Gandhi**[18], **Efrain Garcia**[92], **Donald**

[†]To whom correspondence should be addressed. bwalker@partners.org (B.D.W.); pdebakker@rics.bwh.harvard.edu (P.I.W.d.B.).
[‡]These authors contributed equally to this work.
[*]All authors with their contributions and affiliations appear at the end of this paper.

Supporting Online Material
www.sciencemag.org/cgi/content/full/science.1195271/DC1

Garmon[93], **Joseph C. Gathe Jr.**[94], **Cyril R. Gaultier**[95], **Wondwoosen Gebre**[96], **Frank D. Gilman**[97], **Ian Gilson**[98], **Paul A. Goepfert**[99], **Michael S. Gottlieb**[100], **Claudia Goulston**[101], **Richard K. Groger**[102], **T. Douglas Gurley**[103], **Stuart Haber**[104], **Robin Hardwicke**[105], **W. David Hardy**[24], **P. Richard Harrigan**[106], **Trevor N. Hawkins**[107], **Sonya Heath**[99], **Frederick M. Hecht**[16], **W. Keith Henry**[108], **Melissa Hladek**[109], **Robert P. Hoffman**[110], **James M. Horton**[111], **Ricky K. Hsu**[112], **Gregory D. Huhn**[113], **Peter Hunt**[16], **Mark J. Hupert**[36], **Mark L. Illeman**[114], **Hans Jaeger**[115], **Robert M. Jellinger**[116], **Mina John**[117], **Jennifer A. Johnson**[2], **Kristin L. Johnson**[18], **Heather Johnson**[36], **Kay Johnson**[118], **Jennifer Joly**[64], **Wilbert C. Jordan**[119], **Carol A. Kauffman**[120], **Homayoon Khanlou**[121], **Robert K. Killian**[122], **Arthur Y. Kim**[18], **David D. Kim**[123], **Clifford A. Kinder**[124], **Jeffrey T. Kirchner**[125], **Laura Kogelman**[126], **Erna Milunka Kojic**[127], **P. Todd Korthuis**[128], **Wayne Kurisu**[97], **Douglas S. Kwon**[1], **Melissa LaMar**[93], **Harry Lampiris**[16], **Massimiliano Lanzafame**[129], **Michael M. Lederman**[130], **David M. Lee**[28], **Jean M.L. Lee**[73], **Marah J. Lee**[131], **Edward T.Y. Lee**[132], **Janice Lemoine**[133], **Jay A. Levy**[16], **Josep M. Llibre**[134], **Michael A. Liguori**[112], **Susan J. Little**[22], **Anne Y. Liu**[2], **Alvaro J. Lopez**[135], **Mono R. Loutfy**[136], **Dawn Loy**[137], **Debbie Y. Mohammed**[30], **Alan Man**[35], **Michael K. Mansour**[18], **Vincent C. Marconi**[138], **Martin Markowitz**[139], **Rui Marques**[140], **Jeffrey N. Martin**[16], **Harold L. Martin Jr.**[141], **Kenneth Hugh Mayer**[66], **M. Juliana McElrath**[73], **Theresa A. McGhee**[142], **Barbara H. McGovern**[126], **Katherine McGowan**[2], **Dawn McIntyre**[59], **Gavin X. Mcleod**[143], **Prema Menezes**[81], **Greg Mesa**[144], **Craig E. Metroka**[29], **Dirk Meyer-Olson**[145], **Andy O. Miller**[146], **Kate Montgomery**[147], **Karam C. Mounzer**[148], **Ellen H. Nagami**[1], **Iris Nagin**[149], **Ronald G. Nahass**[150], **Margret O. Nelson**[18], **Craig Nielsen**[151], **David L. Norene**[152], **David H. O'Connor**[153], **Bisola O. Ojikutu**[18], **Jason Okulicz**[154], **Olakunle O. Oladehin**[18], **Edward C. Oldfield III**[155], **Susan A. Olender**[156], **Mario Ostrowski**[136], **William F. Owen Jr.**[157], **Eunice Pae**[1], **Jeffrey Parsonnet**[158], **Andrew M. Pavlatos**[159], **Aaron M. Perlmutter**[160], **Michael N. Pierce**[218], **Jonathan M. Pincus**[161], **Leandro Pisani**[162], **Lawrence Jay Price**[163], **Laurie Proia**[164], **Richard C. Prokesch**[137], **Heather Calderon Pujet**[165], **Moti Ramgopal**[166], **Almas Rathod**[1], **Michael Rausch**[167], **J. Ravishankar**[168], **Frank S. Rhame**[169], **Constance Shamuyarira Richards**[170], **Douglas D. Richman**[22], **Gregory K. Robbins**[18], **Berta Rodes**[171], **Milagros Rodriguez**[162], **Richard C. Rose III**[172], **Eric S. Rosenberg**[18], **Daniel Rosenthal**[173], **Polly E. Ross**[174], **David S. Rubin**[175], **Elease Rumbaugh**[35], **Luis Saenz**[162], **Michelle R. Salvaggio**[176], **William C. Sanchez**[177], **Veeraf M. Sanjana**[178], **Steven Santiago**[162], **Wolfgang Schmidt**[179], **Hanneke Schuitemaker**[180], **Philip M. Sestak**[181], **Peter Shalit**[182], **William Shay**[104], **Vivian N. Shirvani**[183], **Vanessa I. Silebi**[184], **James M. Sizemore Jr.**[185], **Paul R. Skolnik**[89], **Marcia Sokol-Anderson**[186], **James M. Sosman**[153], **Paul Stabile**[187], **Jack T. Stapleton**[188], **Sheree Starrett**[189], **Francine Stein**[83], **Hans-Jurgen Stellbrink**[190], **F. Lisa Sterman**[191], **Valerie E. Stone**[18], **David R. Stone**[192], **Giuseppe Tambussi**[193], **Randy A. Taplitz**[22], **Ellen M. Tedaldi**[194], **Amalio Telenti**[6], **William Theisen**[2], **Richard Torres**[195], **Lorraine Tosiello**[196], **Cecile Tremblay**[197], **Marc A. Tribble**[198], **Phuong D. Trinh**[199], **Alice Tsao**[1], **Peggy Ueda**[1], **Anthony Vaccaro**[200], **Emilia Valadas**[201], **Thanes J. Vanig**[202], **Isabel Vecino**[203], **Vilma M. Vega**[137], **Wenoah Veikley**[107], **Barbara H. Wade**[204], **Charles Walworth**[65], **Chingchai Wanidworanun**[205], **Douglas J. Ward**[206], **Daniel A. Warner**[207], **Robert D. Weber**[208], **Duncan Webster**[209], **Steve Weis**[203], **David A. Wheeler**[210], **David J. White**[211], **Ed Wilkins**[212], **Alan Winston**[84], **Clifford G. Wlodaver**[213], **Angelique van't Wout**[180], **David P. Wright**[214], **Otto O. Yang**[24], **David L. Yurdin**[215], **Brandon W. Zabukovic**[216], **Kimon C. Zachary**[18], **Beth Zeeman**[1], and **Meng Zhao**[217]

[1]Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology (MIT) and Harvard, Boston, MA, USA [2]Department of Medicine, Division of Infectious Disease, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA [3]Harvard-MIT Division of Health Sciences and Technology, Boston, MA, USA [4]Broad Institute of Harvard and MIT, Cambridge, MA, USA [5]Department of Medicine, Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA [6]Institute of Microbiology, University of

Lausanne, Lausanne, Switzerland [7]Department of Medical Genetics, Division of Biomedical Genetics, University Medical Center Utrecht, Netherlands [8]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Netherlands [9]Howard Hughes Medical Institute, Chevy Chase, MD, USA [10]Department of Medicine, Center for Human Genetic Research, MGH, Harvard Medical School, Boston, MA, USA [11]Cancer and Inflammation Program, Laboratory of Experimental Immunology, SAIC-Frederick, NCI-Frederick, Frederick, MD, USA [12]Microsoft Research, Redmond, WA, USA [13]Microsoft Research, Los Angeles, CA, USA [14]Genentech, South San Francisco, CA, USA [15]Department of Medicine, Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA [16]University of California San Francisco, San Francisco, CA, USA [17]Vanderbilt University School of Medicine, Nashville, TN, USA [18]MGH, Harvard Medical School, Boston, MA, USA [19]Stanford University, Palo Alto, CA, USA [20]Weill Medical College of Cornell University, New York, NY, USA [21]Hawaii Center for AIDS, John A. Burns School of Medicine, University of Hawaii, Honolulu, HI, USA [22]University of California San Diego, San Diego, CA, USA [23]University of Pittsburgh, Pittsburgh, PA, USA [24]University of California Los Angeles, Los Angeles, CA, USA [25]Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA [26]Infectious Disease Clinical Research Program, Uniformed Services University of the Health Sciences, Bethesda, MD, USA [27]Summa Health System, Akron, OH, USA [28]Uptown Physicians Group, Dallas, TX, USA [29]St. Luke's Roosevelt Hospital, New York, NY, USA [30]New Jersey Medical School, University Hospital, Newark, NJ, USA [31]Infectious Disease Physicians, Annandale, VA, USA [32]Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, NY, USA [33]St. Leonards Medical Centre, St. Leonards, Australia [34]Ysbyty Gwynedd Hospital, Gwynedd, UK [35]Kaiser Permanente, Portland, OR, USA [36]Tarrant County Infectious Disease Associates, Fort Worth, TX, USA [37]Private Practice of William Barrie, Toronto, Canada [38]Rowan Tree Medical, Wilton Manors, FL, USA [39]Chelsea and Westminster Hospital, St. Stephen's Centre, London, UK [40]Beaver Street Family Practice, Flagstaff, AZ, USA [41]Southwest Infectious Disease Associates, Dallas, TX, USA [42]St. Barnabas Hospital, Bronx, NY, USA [43]Research Institute, McGill University Health Centre, Montreal General Hospital, Montreal, Canada [44]Thacker, Thompson and Bernard, Atlanta, GA, USA [45]Vanderbilt University School of Medicine, Hermitage, TN, USA [46]LA Gay and Lesbian Center, Los Angeles, CA, USA [47]Louisiana State University Health Sciences Center, University Medical Center East Clinic, Lafayatte, LA, USA [48]Physicians' Research Network, Callen-Lorde Community Health Center, New York, NY, USA [49]Brndjar Medical Associates, Allentown, PA, USA [50]AIDS Research Alliance, Los Angeles, CA, USA [51]David Powell Community Health Center, Austin, TX, USA [52]James J. Peters VA Medical Center, Bronx, NY, USA [53]Sunrise Medical Group, Brooklyn, NY, USA [54]University of Miami-Miller School of Medicine, Lake Worth, FL, USA [55]WellSpring Medical Group, San Francisco, CA, USA [56]Moses Cone Health System, Greensboro, NC, USA [57]HealthPartners Infectious Disease, St Paul, MN, USA [58]El Rio Special Immunology Associates, Tuscon, AZ, USA [59]Jersey Shore University Medical Center, Neptune, NJ, USA [60]St. Michaels Hospital, Toronto, Canada [61]The Brooklyn Hospital Center, PATH Center, Brooklyn, NY, USA [62]University of Otago, Christchurch, New Zealand [63]H.E.L.P./Project Samaritan, Bronx, NY, USA [64]David E. Rogers Center for HIV/AIDS Care, Southampton, NY, USA [65]Center for Special Immunology, Fountain Valley, CA, USA [66]Fenway Community Health, Boston, MA, USA [67]9th Street Internal Medicine Associates, Philadelphia, PA, USA [68]University of British Columbia, Vancouver, Canada [69]National Centre in HIV Epidemiology and Clinical Research, Sydney, Australia [70]Metro Infectious Disease Consultants, Indianapolis, IN, USA [71]John H. Stroger Hospital of Cook County, Chicago, IL, USA [72]New England Quality Care Alliance, Braintree, MA, USA [73]Fred Hutchinson Cancer Research Center, Seattle, WA, USA [74]Desert AIDS Project, Palm Springs, CA, USA [75]University of Massachusetts Memorial Medical Center, Worcester, MA, USA [76]University of Iowa Hospitals and Clinics, Iowa City, IA, USA [77]Orlando Immunology Center, Orlando, FL, USA [78]The Kansas City Free Health Clinic, Kansas City, MO, USA [79]Private Practice of Michael E. Dunn, M.D., Tampa,

FL, USA [80]South Shore Hospital, Weymouth, MA, USA [81]University of North Carolina at Chapel Hill, Chapel Hill, NC, USA [82]AIDS Resource Center of Wisconsin, Milwaukee, WI, USA [83]Visiting Nurse Association of Central New Jersey, Community Health Center, Asbury Park, NJ, USA [84]Imperial College, London, UK [85]Heartland Clinic, Paducah, KY, USA [86]Morristown Memorial Hospital, Morristown, NJ, USA [87]Private Practice of Neel K. French, M.D., Chicago, IL, USA [88]San Francisco Department of Public Health, San Francisco, CA, USA [89]Boston University Medical Center, Boston, MA, USA [90]Baystate Medical Center, Springfield, MA, USA [91]Johns Hopkins University School of Medicine, Baltimore, MD, USA [92]Private Practice of Efrain Garcia, M.D., Miami, FL, USA [93]The Rockefeller University, New York, NY, USA [94]Private Practice of Joseph C. Gathe Jr., M.D., Houston, TX, USA [95]Tower Infectious Disease, Los Angeles, CA, USA [96]Nassau University Medical Center, East Meadow, NY, USA [97]Sharp Rees Stealy Medical Center, San Diego, CA, USA [98]Medical College of Wisconsin, Milwaukee, WI, USA [99]University of Alabama, Birmingham, Birmingham, AL, USA [100]Synergy Hematology and Oncology, Los Angeles, CA, USA [101]University of Utah, Salt Lake City, UT, USA [102]South Dayton Acute Care Consultants, Dayton, OH, USA [103]T. Douglas Gurley, M.D., Atlanta, GA, USA [104]St. Vincent's Hospital, New York, NY, USA [105]University of Texas Health Science Center, Houston, TX, USA [106]British Columbia Centre for Excellence in HIV/AIDS, Vancouver, Canada [107]Southwest C.A.R.E. Center, Santa Fe, NM, USA [108]Hennepin County Medical Center, Minneapolis, MN, USA [109]The Catholic University of America, School of Nursing, Washington, DC, USA [110]Mercy Medical Center, Springfield, MA, USA [111]CMC Myers Park Medical Center, Charlotte, NC, USA [112]New York University Medical Center, New York, NY, USA [113]The Ruth M. Rothshon Care Center, Chicago, IL, USA [114]Feldman Medical Group, San Francisco, CA, USA [115]HIV Research and Clinical Care Centre, Munich, Germany [116]Albany Medical College, Albany, NY, USA [117]Murdoch University, Murdoch, Australia [118]University of Cincinnati, Cincinnati, OH, USA [119]OASIS Clinic, Los Angeles, CA, USA [120]VA Ann Arbor Healthcare System, Ann Arbor, MI, USA [121]AIDS Healthcare Foundation, Los Angeles, CA, USA [122]Capitol Hill Medical, Seattle, WA, USA [123]Astor Medical Group, New York, NY, USA [124]The Kinder Medical Group, Miami, FL, USA [125]Lancaster General Hospital, Lancaster, PA, USA [126]Tufts Medical Center, Boston, MA, USA [127]Alpert Medical School of Brown University, Providence, RI, USA [128]Oregon Health and Science University, Portland, OR, USA [129]G.B. Rossi Hospital, Verona, Italy [130]Case Western Reserve University, Cleveland, OH, USA [131]LifeWay, Fort Lauderdale, FL, USA [132]Saint Claire Medical Associate, Toronto, Canada [133]Greater Lawrence Family Health Center, Lawrence, MA, USA [134]Hospital Universitari Germans Trias i Pujol, Barcelona, Spain [135]Infectious Disease Consultants, Tucker, GA, USA [136]University of Toronto, Toronto, Canada [137]Infectious Disease Associates, Sarasota, FL, USA [138]Emory University, School of Medicine, Atlanta, GA, USA [139]Aaron Diamond AIDS Research Center, Rockefeller University, New York, NY, USA [140]Deruico Doencas Infecciosas, Porto, Portugal [141]Park Nicollet Clinic, St. Louis, MN, USA [142]Absolute Care, Atlanta, GA, USA [143]College of Physicians and Surgeons, Columbia University, New York, NY, USA [144]Highland Medical Associates, Hendersonville, NC, USA [145]Medizinische Hochschule, Abteilung Klinische Immunologie, Hannover, Germany [146]Hospital for Special Surgery, New York, NY, USA [147]Family Practice Specialists,Phoenix,AZ,USA [148]Philadelphia FIGHT, Philadelphia, PA, USA [149]Lower East Side Service Center, New York, NY, USA [150]Infectious Diseases Care, Hillsborough, NJ, USA [151]University of Colorado, Denver, Aurora, CO, USA [152]Sutter Medical Group, Sacramento, CA, USA [153]University of Wisconsin in Madison, Madison, WI, USA [154]Brooke Army Medical Center, San Antonio, TX, USA [155]Eastern Virginia Medical School, Norfolk, VA, USA [156]Columbia University Medical Center, New York, NY, USA [157]CA Pacific Medical Center, San Francisco, CA, USA [158]Dartmouth-Hitchcock Medical Center, Lebanon, NH, USA [159]St. Joseph Hospital, Chicago, IL, USA [160]Aaron M. Perlmutter, M.D., Beverly Hills, CA, USA [161]Codman Square Health Center, Dorchester, MA, USA [162]CARE Resource, Miami, FL, USA [163]Castro-Mission Health Center, San Francisco, CA, USA [164]Rush Medical College, Chicago, IL, USA [165]Boulder Community Hospital, Boulder, CO, USA [166]Midway Immunology and

Research Center, Fort Pierce, FL, USA [167]Aerztezentrum Nollendorfplatz, Berlin, Germany [168]State University of New York Downstate Medical Center, Brooklyn, NY, USA [169]Clinic 42, Minneapolis, MN, USA [170]King Edward Memorial Hospital, Paget, Bermuda [171]Fundacion para la Investigacion Biomedica del Hospital Carlos III, Madrid, Spain [172]Summit Medical Group, Knoxville, TN, USA [173]Medical Consultants of South Florida, Coral Springs, FL, USA [174]Western North Carolina Community Health Services, Asheville, NC, USA [175]New York Hospital Medical Center of Queens, Flushing, NY, USA [176]University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA [177]Georgetown University Medical Center, Washington, DC, USA [178]Village Care Health Center, New York, NY, USA [179]Aerzteforum Seestrasse, Berlin, Germany [180]Academic Medical Center Amsterdam, Amsterdam, Netherlands [181]St. Paul's Hospital, Vancouver, Canada [182]Swedish Medical Center, Seattle, WA, USA [183]Cedars-Sinai Medical Center, Los Angeles, CA, USA [184]Mercy Hospital, Miami, FL, USA [185]University of Tennessee, Chattanooga, TN, USA [186]St. Louis University, St Louis, MO, USA [187]William F. Ryan Community Health Center, New York, NY, USA [188]The University of Iowa, Iowa City, IA, USA [189]Rivington House and Village Care, New York, NY, USA [190]Infektionsmedizinisches Centrum Hamburg, Hamburg, Germany [191]California Pacific Medical Center, San Francisco, CA, USA [192]Lemuel Shattuck Hospital, Boston, MA, USA [193]Fondazione San Raffaele Del Monte Tabor, Milan, Italy [194]Temple University School of Medicine, Philadelphia, PA, USA [195]Yale University School of Medicine, Bridgeport, CT, USA [196]Jersey City Medical Center, Jersey City, NJ, USA [197]University of Montreal, Montreal, Canada [198]Baylor University Medical Center, Dallas, TX, USA [199]Montgomery Infectious Disease Associates, Silver Spring, MD, USA [200]Northwestern University, Chicago, IL, USA [201]Hospital de Santa Maria, Faculdade de Medicina de Lisboa, Lisbon, Portugal [202]Spectrum Medical Group, Phoenix, AZ, USA [203]University of North Texas Health Science Center, Fort Worth, TX, USA [204]Infectious Diseases Associates of Northwest Florida, Pensacola, FL, USA [205]Private Practice of Chingchai Wanidworanun, M.D., Arlington, VA, USA [206]Dupont Circle Physician's Group, Washington, DC, USA [207]Consultive Medicine, Daytona Beach, FL, USA [208]Infectious Disease Specialists, Colorado Springs, CO, USA [209]Saint John Regional Hospital, Saint John, Canada [210]Clinical Alliance for Research and Education-Infectious Diseases, Annandale, VA, USA [211]Hawthorn House, Birmingham Heartlands Hospital, Birmingham, UK [212]North Manchester General Hospital, Manchester, UK [213]Private Practice of Clifford Wlodaver, Midwest City, OK, USA [214]Central Texas Clinical Research, Austin, TX, USA [215]Primary Health Care, Des Moines, IA, USA [216]Memorial Neighborhood Health Center Central Clinic, South Bend, IN, USA [217]United Health Services Hospitals, Binghamton, NY, USA [218]All Med and Rehabilitation of New York, Bronx, NY, USA

Abstract

Infectious and inflammatory diseases have repeatedly shown strong genetic associations within the major histocompatibility complex (MHC); however, the basis for these associations remains elusive. To define host genetic effects on the outcome of a chronic viral infection, we performed genome-wide association analysis in a multiethnic cohort of HIV-1 controllers and progressors, and we analyzed the effects of individual amino acids within the classical human leukocyte antigen (HLA) proteins. We identified >300 genome-wide significant single-nucleotide polymorphisms (SNPs) within the MHC and none elsewhere. Specific amino acids in the *HLA-B* peptide binding groove, as well as an independent *HLA-C* effect, explain the SNP associations and reconcile both protective and risk *HLA* alleles. These results implicate the nature of the HLA–viral peptide interaction as the major factor modulating durable control of HIV infection.

Hiv infection is characterized by acute viremia, often in excess of 5 million viral particles per milliliter of plasma, followed by an average 100-fold or greater decline to a relatively stable plasma virus load set point (1). In the absence of antiretroviral therapy, the level of

viremia is associated with the rate of $CD4^+$ T cell decline and progression to AIDS. There is substantial interperson variability in the virus load set point, with most individuals having stable levels exceeding 10,000 RNA copies/ml. Yet a small number of people demonstrate sustained ability to control HIV replication without therapy. Such individuals, referred to as HIV controllers, typically maintain stable $CD4^+$ cell counts, do not develop clinical disease, and are less likely to transmit HIV to others (2).

To determine the genetic basis for this rare phenomenon, we established a multinational consortium (www.hivcontrollers.org) to recruit HIV-1 controllers, who are defined by at least three measurements of plasma virus load (VL) < 2000 RNA copies/ml over at least a 12-month period in the absence of antiviral therapy. We performed a genome-wide association study (GWAS) in the HIV controllers (median VL, CD4 count, and disease duration of 241 copies/ml, 699 cells/mm^3, and 10 years, respectively) and treatment-naïve chronically infected individuals with advanced disease (median VL and CD4 count of 61,698 copies/ml and 224 cells/mm^3, respectively) enrolled in antiviral treatment studies led by the AIDS Clinical Trials Group. After quality control and imputation on the basis of HapMap Phase 3 (3), we obtained data on 1,384,048 single-nucleotide polymorphisms (SNPs) in 974 controllers (cases) and 2648 progressors (controls) from multiple populations (table S1).

After stratification into European, African American, and Hispanic ethnic groups (fig. S1), we tested each SNP for association using logistic regression, including the major principal components as covariates to correct for population substructure (4). In the largest group, comprising 1712 individuals of European ancestry, we identified 313 SNPs with genome-wide significance, defined by $P < 5 \times 10^{-8}$ due to correction for multiple comparisons (table S2). All SNPs that reached genome-wide significance were located in the major histocompatibility complex (MHC) region on chromosome 6 (Fig. 1A). We obtained similar results for the other two ethnic groups and in a meta-analysis of all participants (fig. S2). We also performed a genome-wide analysis to test the influence of local chromosomal ancestry in the African American sample (4), but we detected no signal outside the MHC (figs. S3 and S4). The impact of the MHC was further underscored when we specifically tested published associations related to HIV disease progression outside the MHC. Only variants in the *CCR5-CCR2* locus—namely, CCR5Δ32 deletion polymorphism (5), *C927T* in *CCR5* (6), and $Val^{64} \rightarrow Ile^{64}$ in *CCR2* (7)—replicate with nominal statistical significance in our study (Fig. 1B and table S3).

Closer examination of the significant SNPs within the MHC showed that they are located within a 3-Mb region concentrated around class I *human leukocyte antigen (HLA)* genes (fig. S5), but extensive linkage disequilibrium (LD) makes precise assignment of causal variants challenging (8). Therefore, we used stepwise regression to define independent markers associated with host control. From the initial set of 313 SNPs that reached genome-wide significance in the European sample, for which the greatest numbers of participants were available, we found only four independent markers of association (Table 1). rs9264942, located 35 kb upstream of *HLA-C* and a putative variant associated with *HLA-C* expression levels [odds ratio (OR) = 2.9, $P = 2.8 \times 10^{-35}$, where an OR > 1 indicates a protective effect], and rs2395029, a proxy for *HLA-B*57:01* (OR = 5.3, $P = 9.7 \times 10^{-26}$), had been previously reported to be associated with virus load set point after acute infection (9). We also defined rs4418214, a noncoding SNP near *MICA* (OR = 4.4, $P = 1.4 \times 10^{-34}$), and rs3131018 in *PSORS1C3*, a gene implicated in psoriasis (OR = 2.1, $P = 4.2 \times 10^{-16}$). These four SNPs explain 19% of the observed variance of host control in the European sample; together with those in *CCR5*, these SNPs explain 23%, using Nagelkerke's approximation (Fig. 1C) (10).

In the smaller African American sample, we observed 33 SNPs with genome-wide significance, four of which were identified as independent markers, but all differed from those in the European sample (Table 1). This suggests that shared causal variants are tagged by different SNPs in these two populations or that the mechanism of control differs with ethnicity. Only rs2523608 was previously identified, in a recent study of virus load set point in African Americans (11). Despite no evidence for historical recombination ($D' = 1$), this SNP is only weakly correlated ($r^2 < 0.1$) with *HLA-B*57:03*, the class I allele most strongly associated with durable control of HIV in populations of African ancestry (11-13). In the Hispanic sample, which was much smaller, the most significant SNP was rs2523590, 2 kb upstream of *HLA-B*, also identified in the African American sample described here.

Given the localization of significant SNPs entirely to the HLA class I region, as well as previous studies showing *HLA* alleles to affect disease progression (13-20), we next sought to evaluate whether these SNP and HLA associations might be due to specific amino acids within HLA. Because *HLA* types were available for only a portion of the entire cohort, we developed a method to impute classical *HLA* alleles and their corresponding amino acid sequences (4) on the basis of haplotype patterns in an independent data set collected by the Type 1 Diabetes Genetics Consortium (T1DGC) (21). This data set contains genotype data for 639 SNPs in the MHC that overlap with genotyped SNPs in our GWAS and classical *HLA* types for class I and II loci at four-digit resolution in 2767 unrelated individuals of European descent.

We imputed *HLA* types in the European sample of our study and validated the imputations by comparing to empirical four-digit *HLA* typing data collected for class I loci in a subset ($n = 371$) of the HIV controllers. The quality of the imputations was such that the imputed and true frequencies for all *HLA* alleles in this subset were in near-perfect agreement (Fig. 2A) ($r^2 = 0.99$). Furthermore, the positive predictive value was 95.2% and the sensitivity was 95.2% at two-digit resolution (92.7 and 95.6%, respectively, at four-digit resolution) for *HLA* alleles with frequency >2% (Fig. 2B). This indicates that the performance of the imputation was generally excellent for common alleles, consistent with previous work (22). We used *HLA* allele imputations in all participants (even those with *HLA* types defined by sequencing) for association analyses to avoid systematic bias between cases and controls. Lower imputation quality would only decrease power, not increase the false-positive rate, because cases and controls would be equally affected.

We tested all *HLA* alleles for association via logistic regression, adjusting for the same covariates used in SNP analysis (tables S4 and S5). The most significant HLA association is *B*57:01* (OR = 5.5, $P = 1.4 \times 10^{-26}$), which explains the proxy association of rs2395029 in *HCP5*. With the use of stepwise regression modeling in the European sample of controllers and progressors, we were able to implicate *B*57:01*, *B*27:05*, *B*14/Cw*08:02*, *B*52*, and *A*25* as protective alleles and *B*35* and *Cw*07* as risk alleles. These associations are consistent with earlier studies that highlighted a role for HLA class I loci (13-20), and particularly *HLA-B* alleles in control of HIV, which indicated that the imputations are robust. Collectively explaining 19% of the variance of host control, these *HLA* allele associations are consistent with the effects of the four independent SNPs.

Virus-infected cells are recognized by $CD8^+$ T cells after presentation of short viral peptides within the binding groove of HLA class I, and HIV-specific $CD8^+$ T cells are strongly associated with control (23). We thus evaluated whether the SNP associations identified in the GWAS, and the HLA associations derived from imputation, might be due to specific amino acid positions within the HLA molecules, particularly those involved in the interaction between the viral peptide and the HLA class I molecule. Using the official DNA sequences defined for known *HLA* alleles (24), we encoded all variable amino acid positions

within the coding regions of the *HLA* genes in each of the previously *HLA*-typed 2767 individuals in the T1DGC reference panel, and we used this data set to impute the amino acids in the cases and controls (4). Among a total of 372 polymorphic amino acid positions in class I and II HLA proteins, 286 are biallelic like a typical nonsynonymous coding SNP. The remaining 86 positions accommodate more than two amino acids; position 97 is the most diverse in HLA-B with six possible amino acids observed in European populations.

After imputing these amino acids in the European sample, we used logistic regression to test all positions for association with host control (fig. S6 and table S6). Notably, position 97 in HLA-B was more significant (omnibus $P = 4 \times 10^{-45}$) than any single SNP in the GWAS, and three amino acid positions (67, 70, and 97), all in HLA-B, showed much stronger associations than any single classical *HLA* allele, including *B*57:01* (Fig. 3A). Moreover, allelic variants at these positions were associated with substantial frequency differences between cases and controls (Fig. 3B). These results indicate that the effect of *HLA-B* on disease outcome could be mediated, at least in part, by these positions. These three amino acid positions are located in the peptide binding groove, which suggests that conformational differences in peptide presentation at these sites contribute to the protective or susceptible nature of the various *HLA-B* allotypes. Although both innate and adaptive mechanisms could be at play, the hypothesis that HLA affects peptide presentation and subsequent T cell functionality is supported by experimental data showing substantial functional differences between CTL targeting identical epitopes but restricted by different *HLA* alleles (25).

We next performed stepwise regression modeling and identified six residues as independent markers associated with durable control of HIV. These include Arg^{97}, Cys^{67}, Gly^{62}, and Glu^{63}, all in HLA-B; Ser^{77} in HLA-A; and Met^{304} in HLA-C, which collectively explain 20% of the observed variance (similar to the variance explained by the seven classical *HLA* alleles described above). With the exception of Met^{304} in the transmembrane domain of HLA-C, these residues are all located in the MHC class I peptide binding groove, again suggesting that the binding pocket—and, by inference, the conformational presentation of class I-restricted epitopes—plays a key role in host control.

Having identified these amino acid positions as strong candidates to account for the SNP and HLA association signals in this study, we next investigated their effects on protection or risk, revealing allelic variants at these positions linked to both extremes (Table 2). HLA-B position 97 (omnibus $P = 4 \times 10^{-45}$), located at the base of the C pocket, has important conformational properties for peptide binding (26). Position 97 has six allelic variants: Protective haplotypes *B*57:01*, *B*27:05*, and *B*14* are uniquely defined by Val^{97} (3% frequency in controls), Asn^{97} (4%) and Trp^{97} (3%), respectively; the other amino acids at this position (Ser, Thr, Arg) segregate on a diverse set of haplotypes. Ser^{97} (27% frequency) lies on risk haplotypes *Cw*07*, *B*07*, and others, where-as Thr^{97} (11%) lies on protective *B*52* (and others). Arg^{97} is the most common amino acid (51%) and is carried by risk allele *B*35*, among others. The importance of this amino acid position to host control is underscored by conditional analyses revealing significance when we adjust incrementally for Val^{97} (omnibus test for position 97, $P = 3 \times 10^{-20}$), Asn^{97} ($P = 2 \times 10^{-9}$), and Trp^{97} ($P = 7 \times 10^{-5}$). Thus, at a single position within the peptide binding groove (position 97, C-pocket), discrete amino acids are associated with opposite disease outcomes, even after controlling for *B*57* and *B*27*, alleles associated with host control.

We also found similar discordant associations for alleles at positions 67, 63, and 62 (Table 2), all of which line the α1 helix along the peptide binding groove and help shape the B-pocket (Fig. 4). At position 67 (omnibus $P = 2 \times 10^{-42}$), risk haplotypes *B*35* and *B*07* carry aromatic residues Phe^{67} and Tyr^{67}, respectively, whereas protective *B*57:01*, *B*27:05*, and *B*14* alleles carry sulfur-containing residues Met^{67} or Cys^{67}. Position 62 (P =

5×10^{-27}) is biallelic (Arg/Gly) with the Gly^{62} allele segregating with protective alleles *B*57:01* and *B*58* (<1% frequency, OR = 1.7, $P = 0.2$). Adjacent position 63 ($P = 9 \times 10^{-16}$) is also biallelic (Glu/Asn) with Glu^{63} appearing in complete LD ($D' = 1$) with *B*57:01*, *B*27:05*, and *B*52*. In contrast, at this position the risk alleles *B*07* (14% frequency, OR = 0.5, $P = 1 \times 10^{-7}$) and *B*35* both carry Asn^{63}. Position 70 (omnibus $P = 3 \times 10^{-39}$) accommodates four alleles that are tightly coupled with positions 67 and 97: Ser^{70} appears exclusively with Met^{67} (which defines *B*57* and *B*58*), Gln^{70} with Tyr^{67}, and Lys^{70} with Asn^{97} (*B*27*). Hence, these data create a consistent and parsimonious model that can explain the associations of classical *HLA-B* alleles by specific amino acids lining the binding groove (and residues tightly coupled to them), which are expected to have an impact on the three-dimensional structure of the peptide-MHC complex.

To further investigate the role of individual amino acid positions in *HLA-B*, we implemented a permutation procedure to assess how consistent the above observations are with a null model in which there is no relation between amino acids at a particular position and host control (4). The results of this procedure provided evidence that multiple amino acid positions in the peptide binding groove are indeed associated with host control (table S7), including positions 62, 63, 67, 70, and 97, thus providing a structural basis for the effect of *HLA-B* on host control (Fig. 4).

Within HLA-A position 77, which lies on the α helix contributing to the F-pocket, we identified a weaker but still significant association (omnibus $P = 3 \times 10^{-6}$). Ser^{77} (6% frequency, OR = 2.0, $P = 2 \times 10^{-6}$) is carried by only two *HLA-A* alleles (joint $r^2 = 1$): *A*25* (2.4% frequency, OR = 2.6, $P = 1 \times 10^{-5}$) and *A*32* (3.2%, OR = 1.6, $P = 0.02$). Given its location and earlier association evidence for the A10 supertype (27), *HLA-A* could play a role in host control, although the evidence is not as strong as for *HLA-B*.

The signals within *HLA-C* are less straight-forward to interpret. Position 304 is a biallelic variant (Val/Met) located in the transmembrane domain (Met^{304}, 28% frequency, OR = 2.3, $P = 7 \times 10^{-23}$). Met^{304} is in moderate LD ($r^2 = 0.5$) with rs9264942, which is known to be associated with *HLA-C* expression levels (28). Addition of this SNP to a multivariate model of all six amino acids is marginally significant ($P = 0.013$) but eliminates the effect of Met^{304} ($P = 0.06$). Similarly, addition of rs9264942 to a multivariate model of all seven independent classical *HLA* alleles is also significant ($P = 2 \times 10^{-4}$) but eliminates the effect of *Cw*07* ($P = 0.08$). These observations make it difficult to determine the extent to which epitope presentation in the *HLA-C* peptide binding pocket is important for host control. Thus, rs9264942 could be a proxy for not only many protective and risk *HLA* alleles (predominantly at *HLA-B*), but also for an independent effect on *HLA-C* gene expression, differentially affecting the response to HIV (29).

We next evaluated associations for the SNPs in the MHC, classical *HLA* alleles, and amino acids in a second independent cohort of untreated HIV-infected persons from Switzerland (fig. S7 and tables S8 and S9) (4), in whom virus load set point was measured as a quantitative trait. Allelic variants at positions 67, 70, and 97 were also associated with highly significant differences in virus load set point in this second cohort (Fig. 3C). The effect estimates of all variable amino acids in *HLA-B* ($r^2 > 0.9$) and, to a lesser degree, those in *HLA-C* ($r^2 > 0.8$) in that cohort are in excellent agreement (figs. S8 and S9). As before, position 97 in HLA-B is the most significant association (omnibus $P = 1 \times 10^{-13}$). The HLA-A associations (*A*25* or Ser^{77}) did not replicate, which reduces the likelihood that HLA-A plays a major role in host control.

In the African American sample (fig. S10), the most significant *HLA* allele association was observed for two-digit *B*57* (OR = 5.1, $P = 1.7 \times 10^{-21}$) and four-digit *B*57:03* (OR = 5.1,

$P = 2.8 \times 10^{-17}$; tables S10 and S11), consistent with previous studies (11-13). Position 97 in HLA-B (omnibus $P = 2 \times 10^{-25}$) is again the most significant amino acid (table S12). The consistency of these results demonstrates that imputation and association testing at amino acid resolution in multiple ethnicities can resolve disparate SNP associations in the MHC and help with fine-mapping of classical HLA associations.

Altogether, these results link the major genetic impact of host control of HIV-1 to specific amino acids involved in the presentation of viral peptides on infected cells. Moreover, they reconcile previously reported SNP and HLA associations with host control and lack of control to specific amino acid positions within the MHC class I peptide binding groove. Although variation in the entire HLA protein is involved in the differential response to HIV across *HLA* allotypes, the major genetic effects are condensed to the positions highlighted in this study, indicating a structural basis for the HLA association with disease progression that is probably mediated by the conformation of the peptide within the class I binding groove. The most significant residue, position 97 in the floor of the peptide binding groove of *HLA-B*, is associated with the extremes of viral load, depending on the expressed amino acid. This residue has been shown to have important conformational properties that affect epitope-contacting residues within the binding groove (26, 30) and has also been implicated in HLA protein folding and cell-surface expression (31).

Although the main focus of this study was on common sequence variation, it remains an open question as to the role of variants outside the MHC and the contribution of epistatic effects and epigenetic regulation. Additional factors also contribute to immune control of HIV, including fitness-altering mutations, immuno-regulatory networks, T cell help, thymic selection, and innate effector mechanisms such as killer cell immunoglobulin-like receptor recognition (23), some of which are influenced by the peptide-HLA class I complex. However, the combination and location of the significant amino acids defined here are most consistent with the genetic associations observed being modulated by HLA class I restricted $CD8^+$ T cells. These results implicate the nature of the HLA-viral peptide interaction as the major genetic factor modulating durable control of HIV infection and provide the basis for future studies of the impact of HLA-peptide conformation on immune cell induction and function.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was made possible through a generous donation from the Mark and Lisa Schwartz Foundation and a subsequent award from the Collaboration for AIDS Vaccine Discovery of the Bill and Melinda Gates Foundation. This work was also supported in part by the Harvard University Center for AIDS Research (grant P-30-AI060354); University of California San Francisco (UCSF) Center for AIDS Research (grant P-30 AI27763); UCSF Clinical and Translational Science Institute (grant UL1 RR024131); Center for AIDS Research Network of Integrated Clinical Systems (grant R24 AI067039); and NIH grants AI28568 and AI030914 (B.D.W.); AI087145 and K24AI069994 (S.G.D.); AI069513, AI34835, AI069432, AI069423, AI069477, AI069501, AI069474, AI069428, AI69467, AI069415, Al32782, AI27661, AI25859, AI28568, AI30914, AI069495, AI069471, AI069532, AI069452, AI069450, AI069556, AI069484, AI069472, AI34853, AI069465, AI069511, AI38844, AI069424, AI069434, AI46370, AI68634, AI069502, AI069419, AI068636, and RR024975 (AIDS Clinical Trials Group); and AI077505 and MH071205 (D.W.H.). The Swiss HIV Cohort Study is supported by the Swiss National Science Foundation (SNF grants 33CSC0-108787 and 310000-110012). S. Ripke acknowledges support from NIH/National Institute of Mental Health (grant MH085520). This project has been funded in whole or in part with funds from National Cancer Institute/NIH (grant HHSN261200800001E to M. Carrington). The content of this publication does not necessarily reflect the views or policies of the U.S. Department of Health and Human Services, nor does the mention of trade names, commercial products, or organizations imply endorsement by the U.S. government. This research was supported in part by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

References and Notes

1. McMichael AJ, Borrow P, Tomaras GD, Goonetilleke N, Haynes BF. Nat. Rev. Immunol. 2010; 10:11. [PubMed: 20010788]
2. Deeks SG, Walker BD. Immunity. 2007; 27:406. [PubMed: 17892849]
3. The International HapMap 3 Consortium. Nature. 2010; 467:52. [PubMed: 20811451]
4. See supporting online material on *Science* Online for detailed background on the analyses that we performed.
5. Dean M, et al. Science. 1996; 273:1856. [PubMed: 8791590]
6. Martin MP, et al. Science. 1998; 282:1907. [PubMed: 9836644]
7. Smith MW, et al. Science. 1997; 277:959. [PubMed: 9252328]
8. de Bakker PIW, et al. Nat. Genet. 2006; 38:1166. [PubMed: 16998491]
9. Fellay J, et al. Science. 2007; 317:944. 10.1126/science.1143767. [PubMed: 17641165]
10. Nagelkerke NJD. Biometrika. 1991; 78:691.
11. Pelak K, et al. J. Infect. Dis. 2010; 201:1141. [PubMed: 20205591]
12. Pereyra F, et al. J. Infect. Dis. 2008; 197:563. [PubMed: 18275276]
13. Costello C, et al. AIDS. 1999; 13:1990. [PubMed: 10513667]
14. Klein MR, et al. J. Infect. Dis. 1994; 169:1244. [PubMed: 8195600]
15. Kaslow RA, et al. Nat. Med. 1996; 2:405. [PubMed: 8597949]
16. Carrington M, et al. Science. 1999; 283:1748. [PubMed: 10073943]
17. Migueles SA, et al. Proc. Natl. Acad. Sci. U.S.A. 2000; 97:2709. [PubMed: 10694578]
18. Flores-Villanueva PO, et al. Proc. Natl. Acad. Sci. U.S.A. 2001; 98:5140. [PubMed: 11309482]
19. Carrington M, O'Brien SJ. Annu. Rev. Med. 2003; 54:535. [PubMed: 12525683]
20. Kiepiela P, et al. Nature. 2004; 432:769. [PubMed: 15592417]
21. Brown WM, et al. Diabetes Obes. Metab. 2009; 11(suppl. 1):2. [PubMed: 19143809]
22. Leslie S, Donnelly P, McVean G. Am. J. Hum. Genet. 2008; 82:48. [PubMed: 18179884]
23. Virgin HW, Walker BD. Nature. 2010; 464:224. [PubMed: 20220841]
24. Robinson J, et al. Nucleic Acids Res. 2009; 37:D1013. [PubMed: 18838392]
25. Leslie A, et al. J. Immunol. 2006; 177:4699. [PubMed: 16982909]
26. Fagerberg T, Cerottini JC, Michielin O. J. Mol. Biol. 2006; 356:521. [PubMed: 16368108]
27. Catano G, et al. PLoS ONE. 2008; 3:e3636. [PubMed: 18982067]
28. Stranger BE, et al. Nat. Genet. 2007; 39:1217. [PubMed: 17873874]
29. Thomas R, et al. Nat. Genet. 2009; 41:1290. [PubMed: 19935663]
30. Stewart-Jones GB, et al. J. Immunol. 2005; 175:2459. [PubMed: 16081817]
31. Blanco-Gelaz MA, et al. Int. Immunol. 2006; 18:211. [PubMed: 16361312]
32. Pettersen EF, et al. J. Comput. Chem. 2004; 25:1605. [PubMed: 15264254]
33. Single-letter abbreviations for the amino acid residues referred to in Table 2 are as follows: Glu E, Phe F, Gly G, Lys K, Met M, Asn N, Gln Q, Arg R, Ser S, Val V, Trp W, Tyr Y.

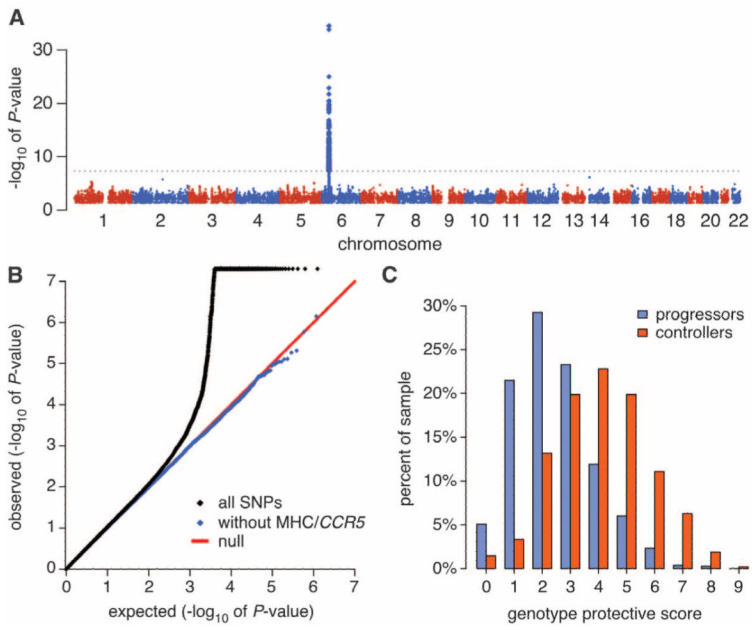

Fig. 1.

Genome-wide association results in the European sample. (**A**) Manhattan plot of 1.3 million autosomal SNPs. Only SNPs in the MHC on chromosome 6 reach genome-wide significance, indicated by the horizontal dotted line ($P < 5 \times 10^{-8}$). Red and blue colors alternate between chromosomes. (**B**) Quantile-quantile plot of the association results with (black) and without (blue) SNPs in the extended MHC and the CCR5-CCR2 locus, indicating that the detectable effect is entirely attributable to these two loci. The red line denotes the expected distribution under the null hypothesis of no effect. (**C**) Distribution of the genotype protective score, defined as the total number of alleles associated with host control at the four independent SNPs in the MHC and the variants at CCR5-CCR2, showing marked differences in controllers (orange) and progressors (blue). In aggregate, these variants explain 23% of the observed variance of durable host control.

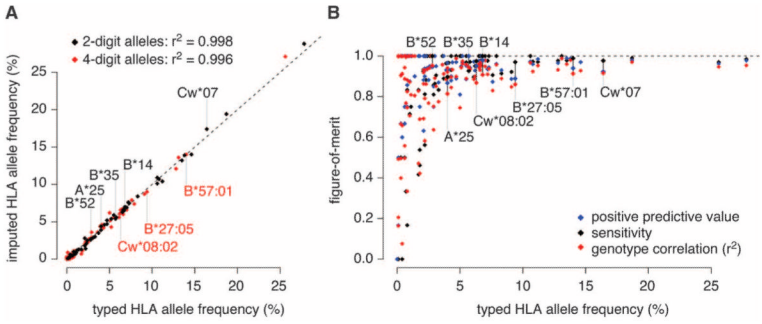

Fig. 2.

Imputation quality of classical *HLA* alleles in the European sample. (**A**) Concordance between imputed (*y*-axis) and observed (*x*-axis) frequencies of classical HLA types in 371 HIV-1 controllers with four-digit *HLA* types obtained through Sanger sequencing. (**B**) Positive predictive value, sensitivity, and genotype correlation (r^2) with typed alleles as a function of the observed frequency.

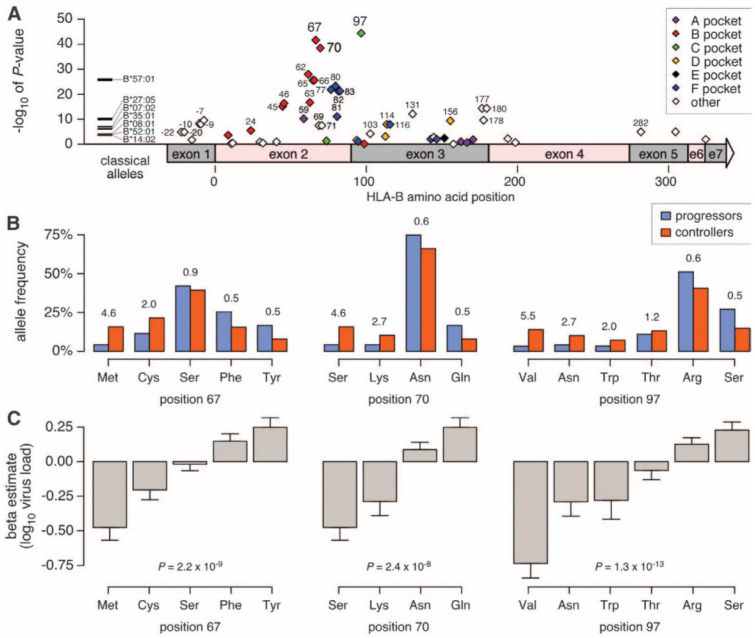

Fig. 3.

Associations at amino acids in HLA-B in the European sample. (**A**) Association results for all variable amino acid positions, as calculated by the omnibus test. Colors denote conventional pocket positions. P values for significant classical HLA-B alleles are shown for comparison. (**B**) Marked allele frequency differences between controllers and progressors for amino acids at positions 67, 70, and 97. Numbers above the bars indicate odds ratios (values >1 indicate a protective effect). (**C**) Associations between allelic variants at amino acid positions 67, 70, and 97 and quantitative virus load set point in the independent Swiss HIV cohort study. Effect estimates (beta coefficients from a linear-regression model) are given in $\log_{10}$ units of virus load set point. P values refer to the omnibus test for association at each position. Error bars indicate the standard error of the beta coefficient.

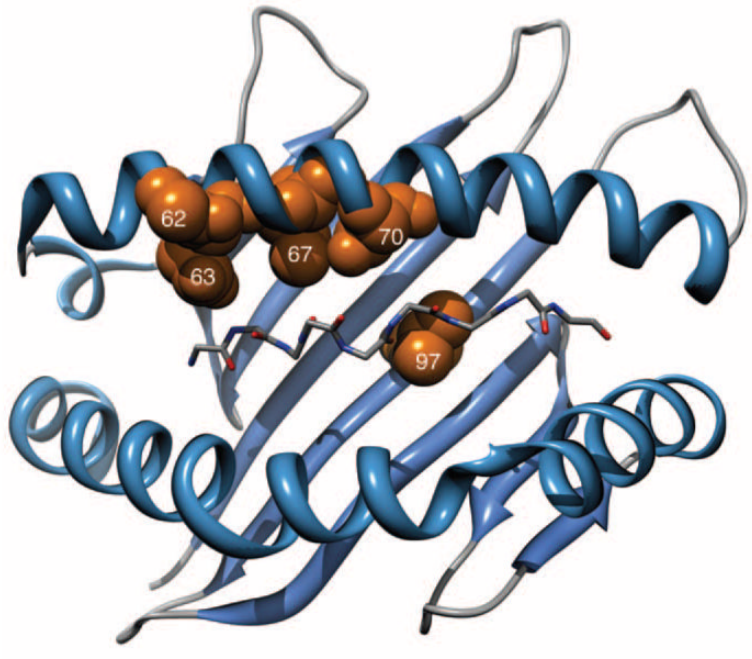

Fig. 4.

Three-dimensional ribbon representation of the HLA-B protein based on Protein Data Bank entry 2bvp (30), highlighting amino acid positions 62, 63, 67, 70, and 97 lining the peptide binding pocket. The peptide backbone of the epitope is also displayed. This figure was prepared with UCSF Chimera (32).

Table 1

Association results for the independent SNPs in the MHC identified with stepwise regression in the European and African American samples. The odds ratio and frequency is given for the *A1* allele, where OR > 1 indicates a protective effect. Odds ratios and *P* values were computed for univariate and multivariate regression models. C, cytosine; G, guanine; T, thymine; A, adenine.

SNP	*A1*	*A2*	Frequency in controllers	Frequency in progressors	Univariate		Multivariate	
					OR	*P* value	OR	*P* value
				European				
rs9264942	C	T	0.595	0.336	2.9	2.8×10^{-35}	2.1	6.3×10^{-16}
rs4418214	C	T	0.240	0.075	4.4	1.4×10^{-34}	1.8	4.9×10^{-4}
rs2395029	G	T	0.139	0.032	5.3	9.7×10^{-26}	2.1	3.5×10^{-4}
rs3131018	C	A	0.777	0.625	2.1	4.2×10^{-16}	1.5	1.2×10^{-5}
				African American				
rs2523608	G	A	0.522	0.326	2.6	8.9×10^{-20}	2.3	3.7×10^{-15}
rs2255221	T	G	0.264	0.137	2.7	3.5×10^{-14}	1.9	2.1×10^{-6}
rs2523590	C	T	0.300	0.164	2.4	1.7×10^{-13}	2.3	1.2×10^{-12}
rs9262632	G	A	0.097	0.034	3.1	1.0×10^{-8}	2.2	2.8×10^{-4}

Table 2

Haplotypes defined by the four independent SNPs, classical *HLA* alleles, and amino acids associated with host control in the European sample. Haplotypes are ordered by the estimated odds ratio, where the most common haplotype was taken as reference (OR = 1). *P* values are for each haplotype tested against all other haplotypes. Only haplotypes with >1% frequency are listed, accounting for >85% of haplotype diversity. *HLA-A* alleles were excluded to limit the number of haplotypes. See (33).

rs3131018	*HLA-C*		rs9264942	*HLA-B*						rs4418214	rs2395029	Frequency	OR	*P* value
	Classical	304		Classical	62	63	67	70	97					
C		M	C	*B**57:01	G	E	M	S	V	C	G	0.060	7.05	1.5E–26
C		M	C	*B**52:01	R	E	S	N	T	T	T	0.011	6.32	4.2E–05
C		V	C	*B**27:05	R	E	C	K	N	C	T	0.051	3.41	1.3E–10
C		M	C		R	E	S	N	T	T	T	0.024	2.78	1.3E–03
C	*Cw**08:02	M	C	*B**14:02	R	N	C	N	W	T	T	0.030	2.58	6.0E–03
C		M	C		R	N	S	N	R	T	T	0.021	2.16	4.2E–02
C		V	C		R	N	F	N	T	T	T	0.021	2.02	4.6E–01
C		M	C		R	N	C	N	R	T	T	0.025	1.58	1.7E–01
C		V	C		R	E	S	N	S	T	T	0.012	1.50	4.5E–01
A		M	C		R	E	S	N	R	T	T	0.067	1.38	6.5E–01
C		M	T		R	N	F	N	T	T	T	0.020	1.29	8.9E–01
A		M	C		R	N	S	N	R	T	T	0.016	1.03	1.7E–01
C		V	T		R	E	S	N	R	T	T	0.168	(reference)	1.6E–03
C		M	C		R	E	S	N	R	T	T	0.022	0.98	4.4E–01
A		V	T		R	E	S	N	R	T	T	0.018	0.87	6.0E–02
C	*Cw**07:01	V	T		R	N	S	N	R	T	T	0.016	0.80	9.5E–02
C	*Cw**07:01	V	T	*B**08:01	R	N	F	N	S	T	T	0.085	0.79	6.0E–05
C		V	T		R	N	Y	Q	T	T	T	0.018	0.67	3.3E–02
A	*Cw**07:02	V	T	*B**07:02	R	N	Y	Q	S	T	T	0.116	0.65	3.2E–08
A		V	T	*B**35:01	R	N	F	N	R	T	T	0.050	0.51	4.3E–06
A		V	T		R	N	F	N	R	T	T	0.017	0.29	4.1E–05